

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28

UNITED STATES DISTRICT COURT  
NORTHERN DISTRICT OF CALIFORNIA  
OAKLAND DIVISION

J. DOE 1, et al.,  
Individual and Representative Plaintiffs,  
v.  
GITHUB, INC., et al.,  
Defendants.

Case Nos. 4:22-cv-06823-JST  
4:22-cv-07074-JST

**CORRECTED PLAINTIFFS'[PROPOSED]  
STIPULATED ORDER REGARDING ESI  
PROTOCOL AND PRODUCTION OF ESI  
AND PAPER DOCUMENTS**

1 To expedite the flow of discovery material and to facilitate the consistency in the format of the  
2 documents to be produced by the Parties in this case, Plaintiffs and Defendants (collectively “Parties”),  
3 by and through their respective counsel, hereby stipulate and agree to the terms of this Stipulated Order  
4 Regarding Protocol for the Production of Documents and Electronically Stored Information (the “ESI  
5 Protocol” or “Protocol”).

6 1. This Protocol supplements all other discovery rules and orders. It streamlines ESI  
7 production to promote a “just, speedy, and inexpensive determination of this action, as required by  
8 Federal Rule of Civil Procedure 1.”

9 2. This Protocol may be modified in the Court’s discretion or by stipulation.

10 3. As in all cases, costs may be shifted for disproportionate ESI production requests pursuant  
11 to Federal Rule of Civil Procedure 26. Likewise, a party’s nonresponsive or dilatory discovery tactics are  
12 cost-shifting considerations.

13 4. A party’s meaningful compliance with this Order and efforts to promote efficiency and  
14 reduce costs will be considered in cost-shifting determinations.

15 5. The parties are expected to comply with the District’s E-Discovery Guidelines and are  
16 encouraged to employ the District’s Checklist for Rule 26(f) Meet and Confer regarding Electronically  
17 Stored Information.

## 18 **I. SCOPE**

19 **A. General.** The procedures and protocols set forth in this ESI Protocol shall govern the  
20 production of relevant electronically stored information (“ESI”) and hard copy documents  
21 in this matter, unless the Parties agree in writing to a change or the Court orders a change.  
22 Nothing in this ESI Protocol shall be interpreted to require production of relevant  
23 information protected by the attorney-client privilege, work-product doctrine, or any other  
24 applicable privilege or other immunity. All Parties preserve all such privileges and  
25 protections, and all Parties reserve the right to object to any such privileges and  
26 protections. The Parties further do not waive any objections to the production,  
27 discoverability, admissibility, or confidentiality of documents.  
28

**B. Cooperation.** The Parties will cooperate on issues relating to discovery of ESI and other documents. As part of this cooperation, the Parties will adhere to guidelines as set forth by the Court and as set forth in the United States District Court Northern District of California's Guidelines for the Discovery of Electronically Stored Information<sup>1</sup> and ESI Checklist for Use During the Rule 26(f) Meet and Confer Process.<sup>2</sup> The Parties recognize that collaborative conduct during discovery reduces litigation costs and delay. The Parties will engage in informal discussions about the discovery of ESI throughout the litigation as necessary. Counsel will be knowledgeable about their respective clients' information technology infrastructure, including how potentially relevant data is stored and retrieved. When appropriate, counsel may engage the assistance of someone knowledgeable about a Party's electronic systems, as necessary.

**C. Transparency.** The Parties are obliged to be forthcoming and transparent in disclosing their use of mechanized tools to cull responsive data and encouraged to bring technically adept personnel together to resolve ESI issues. With the goal of permitting requesting Parties an appropriate level of transparency into a producing Party's electronic search process, without micromanaging how the producing Party meets its discovery obligations and without requiring the disclosure of attorney work product or other privileged information, the Parties will endeavor to be reasonably transparent regarding the universe of documents subject to targeted collections or culling via search terms and other limiters (e.g., date ranges and email domains in metadata fields).

**D. Liaison.**

- **Designation:** The Parties (Plaintiffs, collectively and Defendants, collectively, to the extent possible) will each identify an ESI Liaison, who will be knowledgeable about and responsible for each party's ESI. Any Party is free to change its designated ESI Liaison by providing written notice to the other Parties.

<sup>1</sup> <https://www.cand.uscourts.gov/forms/e-discovery-esi-guidelines/>

<sup>2</sup> [https://www.cand.uscourts.gov/filelibrary/1118/ESI\\_Checklist-12-1-2015.pdf](https://www.cand.uscourts.gov/filelibrary/1118/ESI_Checklist-12-1-2015.pdf)

- **Duties of ESI Liaison:** Each ESI Liaison will be prepared to participate in the resolution of any e-discovery disputes or ESI issues that may arise (or designate another person as primarily responsible) and have access to personnel most knowledgeable about the Party’s electronic systems and capabilities in order to, as appropriate, answer pertinent questions.
- **Time Frame for ESI Dispute Resolution:** Each ESI Liaison will acknowledge receipt of an ESI-related inquiry from another ESI Liaison within 3 business days after the initial inquiry and respond substantively no later than 10 business days after the initial inquiry. If the responding ESI Liaison believes the ESI issue in question is particularly complex and requires more than 10 business days to respond substantively, then within 10 business days the responding ESI Liaison will provide a general explanation of the process necessary to answer the question and provide an estimated response date.

**E. Disputes.** If a dispute remains, the Parties shall meet and confer in good faith on any issue regarding ESI. No Party may seek relief from the Court concerning compliance with the ESI Protocol unless it has conferred in good faith with the affected Parties to resolve any dispute regarding the interpretation or application of any provision of this Protocol or any proposed deviation from the provisions of this Protocol. Any dispute(s) left unresolved after good-faith meet-and-confer efforts may be submitted to the Court for resolution in accordance with the procedure set forth in Section H of the Standing Order for All Civil Cases Before District Judge Jon S. Tigar (the “Discovery Dispute Procedure”).<sup>3</sup>

**F. Protective Order.** All productions are subject to the Stipulated Protective Order (ECF No. 63) (“Protective Order”) entered by the Court in this Action. For the avoidance of doubt, nothing herein shall contradict the Parties’ rights and obligations with respect to any information designated as CONFIDENTIAL, HIGHLY CONFIDENTIAL –

---

<sup>3</sup> <https://www.cand.uscourts.gov/civil-standing-order-jst-7/>

ATTORNEYS EYES' ONLY, or HIGHLY CONFIDENTIAL – SOURCE CODE pursuant to the Protective Order.

## II. DEFINITIONS

- A. “Documents”** shall have the same definition as set forth in the Federal Rule of Civil Procedure 34(a).
- B. “Electronically stored information” or “ESI”** as used herein has the same meaning as in the Federal Rules of Civil Procedure and other applicable law, including Rules 26 and 34(a) of the Federal Rules of Civil Procedure. Examples of ESI include:
- Digital communications (e.g., e-mail, voice mail, text messages, instant messaging, and ephemeral messaging (e.g., SnapChat, etc.);
  - Social media information, including messages, posts, direct messages, and emails;
  - E-mail server stores (e.g., Microsoft Exchange .edb);
  - Word-processed documents (e.g., Word, files and drafts thereof);
  - Spreadsheets and tables (e.g., Excel worksheets);
  - Accounting application data (e.g., QuickBooks);
  - Image and facsimile files (e.g., .PDF, .TIFF, .JPG, .GIF);
  - Sound recordings (e.g., .WAV and .MP3);
  - Video and animation (e.g., .AVI and .MOV);
  - Databases (e.g., Access, Oracle, SAP, SQL data);
  - Contact and relationship management data (e.g., Outlook contacts);
  - Calendar and diary application data (e.g., Outlook calendar entries, blog entries);
  - Presentations (e.g., PowerPoint);
  - Network access and server activity logs;
  - Project management application data;
  - Computer aided design/drawing files;
  - Source code;

- Telemetry Data, including, but not limited to, User Engagement Data and Code Snippets Data;
- API data includes user prompts and completions; and
- Cloud-based or other virtualized ESI, including applications, infrastructure and data.

**C. “Extracted Text”** means the text extracted from ESI, and includes all headers, footers, and document body information when reasonably available.

**D. “Hard Copy Document”** means a document collected from physical files and not from electronic sources.

**E. “Hit Report”** means a report with the number of documents identified (i.e., “hit”) by each keyword search term or Boolean search string.

**F. “Load/Unitization file”** means an electronic file containing information identifying a set of paper-scanned images or processed ESI and indicating where individual pages or files belong together as documents, including attachments, and where each document begins and ends. A Load/Unitization file will also contain data relevant to the individual Documents, including extracted and user-created Metadata.

**G. “Media”** means an object or device, including but not limited to a disc, tape, computer, or other device on which data is or was stored.

**H. “Metadata”** means and refers to application and system information of a file that contains data about the file, as opposed to describing the contents of a file, as the term is described and used in “The Sedona Conference Glossary: E-Discovery and Digital Information Management,” Fifth Ed., February 2020.

**I. “Parties”** means or refers collectively to the named Plaintiffs and Defendants in the above-captioned matter, as well as any later added plaintiffs and defendants. **“Party”** shall refer to a plaintiff or defendant, individually.

**J. “Producing Party”** means or refers to a Party in the above-captioned matter from which production of ESI or Hard Copy Documents are sought.

- 1       **K.**     “**Native,**” “**Native Format,**” or “**Native Data Format**” means and refers to an electronic  
2       document’s associated file structure as defined by the original creating application. For  
3       example, the native format of an Excel workbook is a .xls or .xlsx file.
- 4       **L.**     “**NIST Library**” means the list of over 26 million known, traceable software application  
5       files (i.e., the NIST List or the National Software Reference Library (NSRL)) maintained  
6       by the federal technology agency known as the National Institute of Standards and  
7       Technology (NIST). *See* <https://nvd.nist.gov/ncp/repository>.
- 8       **M.**     “**OCR text**” means text generated through the optical character recognition process of a  
9       Hard Copy Document.
- 10      **N.**     “**Requesting Party**” means or refers to a Party in the above-captioned matter seeking  
11      production of ESI or Hard Copy Documents.
- 12      **O.**     “**Search Terms**” means a search methodology using keywords, terms or phrases including  
13      Boolean-search strings.
- 14      **P.**     “**Static Image(s)**” means a representation of ESI produced by converting a Native file into  
15      a standard image format capable of being viewed and printed on standard computer  
16      systems. A Tagged Image File Format, or “TIFF,” image is an example of a Static Image.  
17      Tagged Image File Format or “TIFF” refers to the CCITT Group IV graphic file format for  
18      storing bit-mapped images of ESI or paper documents.
- 19      **Q.**     “**TAR**” means Technology Assisted Review using computer algorithms or programs to  
20      identify and tag potentially responsive documents based on keywords and other Metadata.
- 21      **R.**     “**Text Messages**” means electronic messages that are sent among users with mobile  
22      devices. Text Messages shall include Short Messaging Service (SMS) and Multimedia  
23      Messaging Service (MMS), as the terms are described and used in “The Sedona  
24      Conference Glossary: E-Discovery and Digital Information Management,” Fifth Ed.,  
25      February 2020.

26   **III.     RULE 26(a) ESI DISCLOSURES**

27

28

1 Pursuant to Federal Rule of Civil Procedure 26(a), and to the extent not already disclosed,  
2 Defendants shall disclose the following information.

3 **A. Custodians.** Each party will provide a list of the ten (10) custodians most likely to have  
4 discoverable ESI in their possession, custody, or control (hereinafter, “Priority Document  
5 Custodians”). The Priority Document Custodians shall be identified by name, title,  
6 connection to the instant litigation, and the type of information under his/her control. The  
7 Priority Document Custodians must include each Defendant’s CEO. As discovery is only  
8 beginning, Plaintiffs do not agree that ten total custodians is an appropriate number and  
9 believe a larger number is very likely necessary in this matter, and therefore Plaintiffs  
10 expressly reserve the right to request inclusion of additional custodians as discovery  
11 progresses. The use of ten Priority Document Custodians is intended to streamline  
12 discovery and produce information from Document Custodians most likely to have  
13 relevant information initially, with additional Document Custodians added as documents  
14 are produced to Plaintiffs that permit an examination of additional individuals likely to  
15 possess relevant information.

16 **B. Additional Custodians:** If, after the Parties identify Priority Document Custodians, a  
17 requesting Party believes that additional document custodians or sources should be added,  
18 then the requesting Party shall advise the producing Party in writing of the proposed  
19 additional document custodians or sources and the basis for the request. If the Parties have  
20 not agreed whether to add the document custodian or source within 30 days of the  
21 requesting Party’s request, then the matter may be brought to the Court in accordance with  
22 the procedures provided herein.

23 **C. Non-custodial Data Sources:** A list of known non-custodial data sources (e.g., shared  
24 drives, servers, etc.), if any, likely to contain discoverable ESI. As part of their non-  
25 custodial-data-sources disclosures pursuant to Rule 26(a), Defendants will identify the  
26 databases that contain structured data, unstructured data, social media posts and other  
27  
28



1 information, training data, source code or algorithms, training manuals, or explanations of  
2 the type of data contained in such structured databases.

3 **D. Data Sources.** For the individual custodians whose Documents and ESI will be searched,  
4 the Producing Party will inform the Requesting Party about custodial data sources and  
5 non-custodial data sources (e.g., shared drives, servers, etc.), that will be searched, and  
6 information sufficient for the Requesting Party to evaluate the reasonableness of each data  
7 source selection. The Producing Party will identify the databases that contain structured  
8 data and unstructured data, including pricing, training data, source code, algorithms  
9 embedded in the structured data, and user/customer lists. With respect to these data  
10 sources, the Producing Party will include any readily available user manuals, training  
11 manuals, or explanations of the type of data contained in structured databases or similar  
12 data sources.

13 **E. Third-Party Data Sources.** The Producing Party will provide the Requesting Party a list  
14 of known third-party data sources, if any, likely to contain discoverable ESI and, for each  
15 such source, the extent to which a Party is (or is not) able to preserve information stored in  
16 third-party data source.

17 **F. Employee Technology Use Policies.** As part of the Parties' initial disclosures, the Parties  
18 will produce a copy of their policy(ies) in effect from 2018 to the present concerning use  
19 by employees of phones, notebook computers, and tablets for business purposes. Such  
20 policies may go by the following names: bring-your-own-device (BYOD) policies,  
21 choose-your-own-device (CYOD) policies, corporate-liable-employee-owned (CLEO)  
22 policies, corporate-owned/personally enabled (COPE) policies, mobile-devise-  
23 management (MDM) policies, and dual-use device policies. If no such policy existed for a  
24 particular or the entire time-period, the Producing Party will notify the Requesting Party in  
25 writing.

26 **G. Preservation of Phone Records & Cell Phones.** The Producing Party will disclose the  
27 steps undertaken to internally preserve phone records and cell phone data for potential  
28

individual custodians. With respect to phone records, the Producing Party will disclose whether notification has been provided to the phone carrier(s) for potential individual custodians' work phone numbers (e.g., office phone, fax number(s)), and cell phone numbers (including cell phones used for both work and personal matters) used from 2018 to the present and/or whether any available internally maintained phone record logs or phone bills have been preserved. The Producing Party will also disclose if they have preserved and/or collected data from any potential custodian's cell phones used from 2018 to the present.

**H. Inaccessible Data.** When certain categories of Documents are not reasonably accessible under Fed. R. Civ. P. 26(b)(2)(C)(i) and are likely to contain discoverable ESI, if the Requesting Party asserts that following production certain ESI is not reasonably usable, they shall be preserved until the Parties reach an agreement on these sources. The Parties shall meet and confer with their respective technology experts to discuss resolving such assertions. If the Parties cannot resolve any such disputes after such a meet and confer has taken place, the issue shall be presented to the Court for resolution.

#### **IV. SEARCH PROCEDURES FOR DOCUMENTS AND ESI**

**A. In General.** Prior to conducting a search for responsive Documents and ESI, regardless of the search methodologies to be employed, the Requesting Party and Producing Party shall meet and confer regarding search methodologies (described below) that the Producing Party proposes to use to identify responsive materials.

**B. Preparation for Meet and Confer.** The Parties agree that negotiations regarding search methodology(ies) and custodians are intended to be a cooperative and interactive process, involving a good-faith exchange of information and proposals to facilitate production of the Documents requested by the Requesting Party (subject to privilege assertions). Prior to the initial meet and confer, the Parties shall be prepared to discuss all items identified Section III, *supra*; and the proposed methodology(ies) for searching and retrieving Documents, as described below:

- **“Search Methodology for ESI”** means the proposed method of searching ESI, including any Search Terms and/or Boolean search strings, date ranges and phrases to be searched, and any additional parameters to be applied for ESI; and
- Whether certain categories of documents or requests for production are amenable to retrieval without utilizing Search Terms, i.e., using a “go-get” approach.
- **“Search Methodology for Hard Copy Documents”** means the proposed methodology and scope for manual retrieval of Hard Copy Documents.

**C. TAR/Advanced Analytics.** The present intention of the Parties is to conduct document review and production without any use of TAR or similar technology-assisted methodologies designed to limit the universe of responsive Documents to review. If a Party should change its intention and resolves to use TAR or similar methodology, the Party shall promptly notify the other Parties and the Parties shall promptly meet and confer regarding any related issues up to and including the negotiation of a stipulation governing TAR methodologies and review of documents if the Parties deem it necessary.

- If a Party proposes the use of TAR, within thirty (30) days of providing the required notice to the Parties it shall disclose sufficient information to allow the other Parties to evaluate the TAR process, including (a) the name of the TAR software and vendor, (b) a general description of how the Producing Party’s TAR process will work, including how it will train the algorithm, such as using exemplars, keyword search strings, or some other method, (c) a general description of the categories or sources of the documents included or excluded from the TAR process, and (d) what quality control measures will be taken (“TAR Disclosures”).
- Within fourteen (14) days of receiving a Producing Party’s TAR Disclosures, the Requesting Party may raise with the Producing Party concerns regarding the proposed TAR process or categories of documents it proposes should be excluded from the TAR process. A Requesting Party may also propose exemplars to be used to train a TAR tool or narrow keyword search strings to be used to generate

1                   exemplars to train a TAR tool. Any disputes regarding TAR will not be submitted  
2                   to the Court, Magistrate Judge, or appointed Special Master before the Parties have  
3                   met and conferred in good faith.

4           **D.     Responsive ESI Discovered Outside of Search Methodology.** Notwithstanding any of  
5           the foregoing, if any Party identifies ESI or Documents that are responsive and subject to  
6           discovery in this matter outside of the search methodology(ies) agreed to by the Parties, it  
7           shall produce all such Documents to the extent they are not privileged.

8           **E.     Procedure for Search Terms.** The Parties recognize that developing efficient Search  
9           Terms is an iterative process and will require transparent and cooperative efforts by both  
10          the Producing and Requesting Parties. The Parties shall abide by the following schedule,  
11          which may be modified upon written agreement of the Parties:

- 12           •       Should the Parties agree to using Search Terms, the Requesting Party shall propose  
13           Search Terms in the first instance.
- 14           •       Within fourteen (14) calendar days of receipt of the proposed Search Terms from  
15           the Requesting Party, the Producing Party and Requesting Party shall meet and  
16           confer to attempt to reach agreement on the Search Terms and/or limiters to be  
17           used to identify responsive Documents. To facilitate such meet and confer efforts,  
18           the Producing Party shall discuss, to the extent not protected by the attorney-client  
19           privilege or work-product doctrine, whether the Search Terms should include
  - 20               ○ relevant terminology, nicknames, code words, euphemisms, acronyms, slang,  
21               terms of art, email addresses, and other language likely to be contained in  
22               responsive documents;
  - 23               ○ appropriate synonyms for proposed Search Terms and variations on search  
24               syntax, for example, using wildcard characters, truncation, stem words, fuzzy  
25               logic and proximity terms to address over- and under inclusiveness; and
  - 26               ○ whether any appropriate testing and quality control measures were conducted  
27               or will be conducted to evaluate the sufficiency of the Search Terms.

- Within fourteen (14) calendar days of the Parties' initial meet and confer regarding the proposed Search Terms as set forth herein, the Producing Party shall provide the Requesting Party with a preliminary report of the number of documents identified by the Search Terms (a "Hit Report") and any proposed modifications to the Search Terms. A Hit Report must be provided in Excel (or other format as agreed to by the Parties) and must include (a) the overall number of documents being searched, (b) number of documents hit by the Search Term with separate hits for individual documents hit vs. documents hit including families, and (c) unique documents hit by search string (i.e., documents that would not be included in the search results but for the inclusion of that particular search string). The Producing Party should provide information sufficient to support its modifications to specific Search Terms, e.g., estimates of the incremental number of false positive hits and the incremental number of true positive hits introduced by the disputed Search Terms, as well as examples of the false positive hits. Within seven (7) calendar days after receipt of the Hit Report and proposed modifications, the Requesting Party shall accept, object, or otherwise supplement the proposed modifications. Within seven (7) calendar days thereafter, the Parties shall meet and confer as to any objections and alterations. If the objections are not resolved, the matter will be promptly submitted to the Court for resolution pursuant to the Discovery Dispute Procedure.
- The Parties shall agree on the meaning of the search syntax used by the Parties in developing the Search Terms. Should the Producing Party learn that the search syntax used by its search tool differs from those used by the Parties during Search Term negotiations and such differences materially alter the nature of the agreed upon Search Terms, the Parties shall promptly meet and confer regarding any modifications to the Search Terms so affected.

1       **F. Production and Supplementation of Searches for ESI.** Production of responsive, non-  
 2 privileged documents shall proceed on a rolling basis and be completed within the time-  
 3 period specified by order of the Court. If, after reviewing ESI and Documents received  
 4 from the Producing Party, the Requesting Party believes that additional search terms are  
 5 necessary or that changes to the search terms or method of searching are warranted, the  
 6 Requesting Party shall request a meet and confer under this Paragraph, and the Parties  
 7 shall meet and confer promptly.

8       **G. Validation.** The Parties shall meet and confer regarding appropriate methods to validate  
 9 the recall of the Producing Party's search methodology for custodial ESI.

10       **H. Good Cause Inability of a Party to Meet Deadlines Imposed in this Order.** It is  
 11 expected that the Parties shall make their best efforts to complete the above steps in a  
 12 reasonable and efficient manner. The Parties understand that technical (or other) issues or  
 13 unanticipated volumes may interfere with a Party's best efforts to comply. Should a Party  
 14 anticipate that for good cause it may be unable to meet a deadline ultimately agreed  
 15 between the Parties or ordered by the Court, that Party shall promptly raise the issue with  
 16 the other Parties, explain the reason for the inability to timely comply, and negotiate a  
 17 reasonable extension for compliance. If the Parties are unable to immediately agree upon a  
 18 revised deadline for compliance, they shall promptly raise the issue with the Court for  
 19 resolution. This provision shall not be construed as blanket permission for a Party to  
 20 modify or extend the ultimate deadlines agreed to by the Parties or ordered by the Court  
 21 without good cause, but rather, to recognize that when dealing with search and review of  
 22 large volumes of ESI, there are sometimes legitimate, unanticipated challenges that may  
 23 interfere with a Party's best efforts to fulfill its obligations and therefore, to afford the  
 24 Parties reasonable flexibility and mutual accommodation should such eventuality occur.

## 25       **V. FORMAT OF PRODUCTION**

26       **A. Hard Copy Document Production Format.** The following provisions shall generally  
 27 govern the production format and procedure for Hard Copy Documents and images.  
 28

- All documents originating in hardcopy format will be produced as black-and-white or color (if originally in color), single-page, 300 dpi Group IV tagged image file format (“TIFF”) images, with OCR text and related path provided in document level text files.
- In scanning hardcopy documents, distinct documents should not be merged into a single record, and single documents should not be split into multiple records (i.e., hardcopy documents should be logically unitized). The Producing Party will use reasonable efforts to unitize documents correctly.
- Where a document, or a document group—such as folder, clipped bundle, or binder—has an identification spine or other label, the information on the label shall be scanned and produced as the first page of the document or grouping.
- Productions of the images shall be made using an image load file (.OPT or .LFP) and a delimited database/Metadata load file (.DAT), pursuant to Exhibit 2. Exhibit 1 part 1 lists the objective coding fields that will be produced for hard copy documents.
- The Producing Party shall utilize best efforts to ensure paper records for a particular custodian included in a single production are produced in consecutive Bates-stamp order.

**B. ESI Production Format.** The following provisions shall generally govern the production format and procedure for ESI and are subject to the other provisions contained herein.

- **Production of electronic documents.** Generally, the Producing Party must produce in native format. The Producing Party must produce all ESI that cannot be produced in native format in TIFF format according to the following specifications.
  - Where technically feasible, TIFFs shall be produced as true color, single-page Group IV TIFF in 8½ X 11-inch page size images at a resolution of at least 300 DPI with the quality setting of 75% or higher.

- When producing documents in TIFF format, the image files shall be produced along with Concordance/Opticon image load files, linking the images to the corresponding document that indicate the beginning and ending of each document, showing the Bates number of each page and the appropriate unitization of the documents.
- Each image file of an electronic document will be created directly from the original electronic document. Image files shall show all text and images that would be visible in the original electronic format (Native Format), including redlines and speaker notes.
- All TIFF files are to be provided with an accompanying searchable text (.TXT) file extracted from the native, electronic file (not generated as an OCR file from the TIFF image(s)), and such text files shall contain the full text extraction. To the extent technically feasible, extracted text shall provide all comments, tracked changes, speaker's notes, and text from hidden worksheets, slides, columns and rows. In the case of files with redacted text, OCR'ed text of the redacted documents may be provided in lieu of extracted text. OCR software should be set to the highest quality setting during processing.
- All documents shall be produced in their original language. For documents in foreign languages, the OCR shall be performed using an OCR tool and settings suitable for the particular byte or multi-byte languages.
- Each text file shall be named according to the Bates number of the first page of the corresponding image files (e.g., BATES000001.TXT).
  - a. Microsoft Word Documents (or similar) (.DOC, .DOCX, or substantially similar non-Microsoft file formats) should be produced as native files for each Document, containing all images for that document, and should not be altered to remove metadata, tracked changes, or comments.



○ To the extent possible, email messages and their attachments shall be produced in their native format, e.g., .eml. If email messages and/or attachments cannot be produced in native format and instead must be imaged, the corresponding text file shall include, where reasonably available: (1) the individual(s) to whom the communication was directed (“To”); (2) the author(s) of the email communication (“From”); (3) who was copied and blind copied on such email (“CC” and “BCC”); (4) the subject line of the email (“RE” or “Subject”); (5) the names of any attachments; (6) the date the message was sent or received (“Sent” or “Received”); (7) the name of anyone tagged in the message body, e.g., @John Doe; (8) the filenames of any attachment to the message (“Attachments”); and (9) the entire text (body) of the email, including prior messages included beneath the message being produced, i.e., the “Chain.”

- **Fields and Metadata.** Exhibit 1 part 2 lists the objective coding fields and Metadata that will be produced for all ESI—including ESI produced in TIFF or Native Format.
- **Native Format Production Documents.**
  - In addition to Word files and emails and their attachments, the Parties shall also produce the following ESI types in native file format:
    - a. Excel files;
    - b. Text message files;
    - c. Social media information, including messages, posts, direct messages, and emails;
    - d. Presentation files (e.g., PowerPoint);
    - e. Databases (MS Access);
    - f. Audio, video, animation, and/or image files;
    - g. Web pages;
    - h. Accounting application data;

- i. Contact and relationship management data;
- j. Calendar and diary application data;
- k. Network access and server activity logs;
- l. Project management application data;
- m. Computer aided design/drawing files;
- n. Source code;
- o. Telemetry data, user engagement data, and code snippets data;
- p. API data, user prompts, and completions; and
- q. Cloud-based or other virtualized ESI.

- To the extent responsive Text Messages are produced, they shall be produced in a reasonably usable format. The Producing Party will disclose its planned Text-Message production format to the Receiving Party prior to the production of Text Messages. The Receiving Party retains its rights to meet and confer on the production format to address any concerns. The Parties reserve the right to request production of other ESI types in Native Format. The Parties agree to meet and confer regarding such requests. Any Native files that are produced should be produced with a link in the NativeLink field, along with all extracted Metadata fields as set forth in Exhibit 1.
- PowerPoint or other presentation files should be produced in Native Format (e.g., as .PPT files). PowerPoint presentations shall also be produced in full-slide image format, along with speaker notes (which should follow the full images of the slides) with related searchable text, Metadata, and bibliographic information.
- In the case of personal database (e.g., MS Access) files containing confidential or privileged information, the parties shall meet and confer to determine the appropriate form of production.

- All social media information, messages, posts, emails, direct messages, and forensic data for any and all social media platforms shall be produced in a usable electronic format that provides a “timeline” and context between images, videos, and messages/posts.
- ESI shall be processed in a manner that preserves hidden columns or rows, hidden text, worksheets, notes, tracked changes, and comments. Any Party seeking a deviation from this provision must provide notice to other Parties and the Parties agree to meet and confer regarding such requests.
- Any Document produced in native format, will be produced according to the following specifications:
  - a. A unique Bates number and confidentiality designation shall be used as the file name and the original file name and file extension shall be preserved in the corresponding load file. An example of this convention would be: “BATES000001\_HighlyConfidential.xls”
  - b. The native format Documents shall be accompanied by reference information that sets forth for each document, sufficient information to allow the Parties to track and authenticate the Native Format documents produced, including: (i) the name of the custodian from whose files the electronic file is produced; (ii) an appropriately calculated “MD-5 Hash Value”; (iii) the original name of the file; and (iv) a Bates number.
  - c. In all cases, unless there is no textual content, an OCR or Extracted Text file shall be produced along with the native file. For any Native Format documents that cannot be imaged or where the image is produced as a separate document, a single page placeholder image shall be provided that indicates the file was produced in native

format and that contains the Bates number and confidential designation, if any, of the corresponding file.

- d. To preserve the integrity of any file produced in Native Format, no Bates number, confidentiality designation, or internal tracking number should be added to the body of the Native Format document unless otherwise agreed to between the Producing Party and the Receiving Party.
- e. The Receiving Party may also request that the Producing Party produce additional file types of electronic documents in Native Format where converted image formats distort information or cause it to be improperly displayed. The Parties shall meet and confer regarding such requests in good faith and cooperate consistent with the provisions set forth in this Protocol.

**C. Deduplication and De-NISTing Documents.** To reduce the unnecessary costs of reviewing and producing exact duplicate documents (i.e., the documents and Metadata fields are identical copies), to the extent reasonably possible and accounting for any technical limitations of any Party's ESI, each Party may remove duplicate ESI prior to producing documents.

- Subject to the provisions in this section, each Party may remove duplicate ESI prior to producing documents (based on the MD5 or SHA-1 hash values at the parent document level). Stand-alone files will de-duplicate against other stand-alone files, but not against attachments contained in document families. Where any exact duplicates have attachments, hash values must be identical for both the document-plus-attachment (including associated Metadata other than custodian) as well as for any attachment (including associated Metadata other than custodian) standing alone. Provided, however, that if an "identical" document is an attachment to different emails (e.g., where attachment 1 is sent to person A one day

and then sent to person B another day), then the attachment shall be produced together with each email.

- For exact duplicate documents, the Producing Party will produce certain Metadata described herein for a single production copy. The Producing Party may deduplicate documents within custodians and/or across custodians. Names of custodians of copies of files eliminated due to global deduplication shall be provided in the Duplicate Custodian field accompanying the original file.
- If during its review the Producing Party identifies a large number of duplicate documents, the Parties may meet and confer regarding a custom deduplication protocol. Except for the removal of extracted logos, no custom deduplication method will be implemented without meeting and conferring with the Requesting Party.
- Common system and program files as defined by the NIST Library need not be processed, reviewed, or produced.

**D. E-mail Thread Suppression:** Email threads are email communications that contain prior or lesser-included email communications. A Producing Party may not use e-mail thread suppression to exclude email from production without first meeting and conferring with the Requesting Party regarding the methodology used for suppression and means of preserving and producing Metadata of any suppressed email. The Producing Party is not required to obtain the agreement of the Requesting Party to its threading methodology. However, the Requesting Party reserves the right to raise issues with the Court if it believes the threading methodology is improper after meeting and conferring.

**E. Password Protected Files.** The Producing Party shall make reasonable efforts to ensure that all encrypted or password-protected Documents are successfully processed for review and production under the requirements of this ESI Protocol, and the decrypted Document is produced (if responsive and not privileged). To the extent encrypted or password-protected Documents are successfully processed according to the requirements of this ESI

1 Protocol, the Producing Party has no duty to identify the prior encrypted status of such  
2 Documents. A document that would otherwise be produced in Native Format may be  
3 imaged and produced as TIFFs if necessary to comply with this provision, but the  
4 encrypted status of such documents must be disclosed. ESI that cannot be reviewed  
5 because proprietary software is necessary to view the ESI will be disclosed to the  
6 Requesting Party. To the extent other Documents are not successfully processed that are  
7 members of a responsive family, the Producing Party agrees to: (a) produce a slipsheet for  
8 each encrypted or password protected Document indicating that the Producing Party could  
9 not successfully process the Document because it cannot be decrypted; and (b) provide the  
10 Metadata for the Document required by Exhibit 1 to the extent it can be reasonably  
11 extracted from the file in its encrypted form. The Requesting Party, Receiving Party, and/or  
12 Producing Party, as appropriate, shall meet and confer regarding any requests that a  
13 Producing Party undertake additional efforts to decrypt files after production.

14 **F. Prior or Other Production.** To the extent a set or sets of Documents in a Party's  
15 possession, custody, or control that were previously produced in a prior or other  
16 proceeding or litigation are the subject of discovery in this case and the ESI for such  
17 Documents (including but not limited to Metadata fields, searchable text, and  
18 organization) differs from the substance and format agreed to herein, such Party may  
19 produce the production set in the same format it was produced in the prior  
20 proceedings/litigation, including, but not limited to, with whatever Metadata fields,  
21 searchable text, organization, and formats that were produced in that litigation.

- 22 • In the event certain Documents in the prior or other productions are in a form that  
23 is not reasonably usable by the Requesting Party, the Requesting Party and the  
24 Producing Party shall meet and confer regarding alternative forms of production  
25 for such Documents. If the Parties are unable to agree after meeting and conferring  
26 in good faith, the Parties may seek resolution by the Court, Magistrate Judge or  
27 Special Master appointed by the Court.

- If materials are being produced from a prior or other proceeding or litigation, the Producing Party shall specify the proceeding or litigation in which each document was previously produced, as well as the entity that produced it. If the prior production contained a unique Bates label, it is sufficient for the Producing Party to identify the prior productions with reference to that Bates label. For example, if Documents being produced in this litigation bear a prior Bates stamp of “ABC-xxxxxx,” the Producing Party can identify the prior proceeding(s) in which such documents were produced by disclosing that all Documents bearing a Bates stamp of “ABC-xxxxxx” reflect documents produced by Party A in litigation Y.
- Notwithstanding the foregoing, any prior production document produced in this case, must, in addition to its prior Bates labeling, also be Bates labeled for this case, as set forth below.

**G. Prior Privilege Logs.** To the extent a Party produces Documents in this case that were produced in a prior proceeding or litigation, the producing Party may produce the privilege logs from those/ prior proceedings/litigations in lieu of creating, in this case, a new log (or log entries) for the Documents that were withheld in the prior proceeding/litigation. The Requesting Party retains the right to request additional information regarding any entry on each such prior log. If the Producing Party objects to producing such additional information, the Requesting Party retains the right to all related challenges, including, without limitation, the right to challenge the sufficiency of the prior log (whether as a whole and/or as to particular entries), and/or the validity of any privilege claim(s) asserted.

**H. Redactions.**

- To the extent that a responsive document contains privileged content, or to comply with federal or state privacy requirements (e.g., HIPAA, or information related to medical or health issues of an individual), the Producing Party may produce that document in a redacted form. Any redactions shall be clearly indicated on each

1 page of the document from which information is redacted, bearing a designation  
2 that it has been redacted. The designation shall make clear the reason for the  
3 redaction (e.g., “Redacted Privileged”). Where a document contains both  
4 privileged and non-privileged responsive content, the Producing Party shall redact  
5 the privileged material and produce the remainder of the document as redacted.

6 ○ The Parties may also redact personal information to the extent that information  
7 is comprised of the following: social security numbers, taxpayer-identification  
8 numbers, driver’s license numbers, passport numbers, financial-account  
9 numbers or other bank account information, or personal passcodes. Any such  
10 redactions in this category should be clearly identified as “Redacted – Personal  
11 Data.”

12 • Native Excel or other spreadsheet files which are redacted may be redacted by  
13 overwriting the data contained in a particular cell, row, column, or tab with the  
14 word “Redacted” and shall make clear the reason for the redaction (e.g., “Redacted  
15 Privilege”). The Parties acknowledge that redacting the native may alter the  
16 Metadata of the produced file; however, the Metadata produced subject to Exhibit  
17 1 will be the Metadata extracted from the original native file, other than Metadata  
18 containing redacted information.

19 • Native Excels or other spreadsheet files, PowerPoint or other presentation files  
20 which are redacted may be produced as black-and-white or color (as applicable),  
21 single-page, 300 DPI Group IV TIFF images, with text and related path provided  
22 in document level text files, with load files pursuant to Exhibit 2, which shall  
23 include all information contained in the original file, including but not limited to  
24 speaker notes and “hidden” slides, with any redaction clearly delineated. If the file  
25 contains external video or audio components, the video or audio will be produced  
26 as native files.



- If the items redacted and partially withheld from production are audio/visual files, the producing party shall, to the extent reasonably practicable, provide the unredacted portions of the content. If the content is a voice recording, the parties shall meet and confer to discuss the appropriate manner for the producing party to produce the unredacted portion of the content.
- **No “Relevancy” Redactions.** The Parties will not make any redactions based upon the purported relevancy of a Document. The only redactions permitted are those on the basis of privilege or the specific categories of medical or personal information identified in Paragraph V.H. herein.

**I. Exception Logs:** ESI that cannot be reviewed, produced and/or imaged because of technical issues should be identified as exception files and included on a log that lists the file name, custodian, and reason for exception: e.g., corruption, unavailable password protection, proprietary software, or other technical issues. The Producing Party shall provide an updated copy of this log to the receiving party at least once a month, and shall provide a final copy of the log upon completion of document production. If the Receiving Party requests production of any files listed on the exception log, the Parties will meet and confer on a reasonable and cost-effective means for attempting to provide the requested files.

**J. Confidentiality of Produced ESI.** Responsive ESI, whether produced as TIFF images or in Native Format, shall be produced pursuant to the terms of the separately entered Protective Order. If the Producing Party is producing ESI in TIFF Format subject to a claim that it is protected from disclosure under any protective agreement or confidentiality order, or any agreement entered into or Order issued in this matter, the word “CONFIDENTIAL” (or something substantially similar) shall be annotated on each page of such document.

- If the Producing Party is producing ESI in Native Format subject to a claim that it is protected from disclosure under the Protective Order or any protective

agreement or confidentiality order entered into or other Order issued in this matter, then the designation shall be included in the filename (to the extent it is technologically feasible) as well as the TIFF placeholder or where not produced with a TIFF placeholder, the storage device (e.g., CD, USB, or hard drive) containing such native ESI data shall be labeled with the designation “CONFIDENTIAL” or “HIGHLY CONFIDENTIAL – ATTORNEYS’ EYES ONLY” in a location that does not obliterate, conceal, or interfere with any information from the source document.

**K. Additional Production Specifications.** All productions will include these additional specifications:

- Delimited Metadata load file (.DAT) containing a field with the full path and file name to the native files produced pursuant to Exhibit 2, and shall include the Metadata fields identified below (Exhibit 1);
- Delimited image load/unitization file for images (.opt, .lfp, or .dii) shall be provided pursuant to Exhibit 2;
- Document level .txt files for all native documents containing extracted full text or OCR text (OCR only if extracted text is not available or if the document has been redacted);
- Bates number branding and Confidentiality designation (if any) on the face of the TIFF image in the lower right-hand corner using a consistent font type and size that, to the extent reasonably feasible, does not obliterate, conceal, or interfere with any information from the source document, including bates number branding from previous litigation productions. If the placement in the lower right-hand corner will result in obscuring the underlying image, the Bates number should be placed as near to that position as possible while preserving the underlying image;
- A TIFF placeholder embossed with the corresponding confidentiality designation and Bates number shall be produced for all ESI produced in Native Format;

- Each Bates number will: (i) be unique across all document productions, including prior productions; (ii) maintain a constant length and format across the entire production (i.e., ABC00000001- with a consistent prefix, no space between the prefix and the number, and padded to the same number of characters); and (iii) contain no special characters; and (iv) be sequential within a given document. If a Bates number or set of Bates numbers is skipped in a production, the Producing Party will so note in a cover letter or production log accompanying the production.
- To the extent any documents contain redactions and are produced in TIFF format, the following specifications apply:
  - PowerPoint slides/presentations that are produced in TIFF format will be produced with any notes, talking points, outlines, or comments contained therein; and
  - All hidden text (e.g., track changes, hidden columns, comments, notes, markups, etc.) will, to the extent reasonably feasible, be expanded, extracted, and rendered in the TIFF file.
- Each of the Metadata and coding fields set forth in Exhibit 1 that can be extracted from a Document shall be produced for that Document. The Parties are not obligated to populate manually any of the fields in Exhibit 1 *except* Default Production Fields (“DPF”) that are generated in the course of collection, review and production (e.g., BEGBATES, ENDBATES, etc.); and the “CUSTODIAN,” “DUPLICATE CUSTODIAN,” and “REDACTION” fields. If Metadata fields are not produced because they do not exist or because they are not reasonably accessible, the Producing Party shall include the fields in the header row of the .DAT and leave the field blank for the corresponding document(s).
- **Time Zone.** ESI items shall be processed in a manner that preserves their existing time, date, and time-zone Metadata (e.g., the email of a Document Custodian

located in Pennsylvania will be processed as Eastern Time, while a Document Custodian located in California will be processed as West Coast Time). If GMT time zone is used, then a time-zone offset Metadata field must be provided indicating the original time zone in which the custodian of the document received or authored the document.

- **Auto date/time stamps.** ESI items shall be processed so as to preserve the date/time shown in the document as it was last saved, not the date of collection or processing.
- The Parties understand and acknowledge that such standardization affects only dynamic date fields and Metadata values and does not affect, among other things, dates and times that are hard-coded text within a file.

## VI. PRESERVATION

**A. Deleted Files and Meet and Confer.** The Parties shall not be obligated under this Protocol to produce or preserve ESI that was deleted, lost, or destroyed before the date upon which the duty to preserve ESI arose, as a result of the routine, good-faith operation of an ESI system. Nothing in this provision limits or expands a Party's obligation to seek or search for ESI known to be "deleted" by a user that remains as a copy on an archive, backup, or as active data on some other central server or place within a Party's possession, custody, or control provided such search is otherwise consistent with and does not exceed or limit the Party's obligations under the Federal Rules of Civil Procedure. If a Producing Party learns that responsive ESI that should have been preserved was lost or destroyed, and is no longer retrievable, the Parties shall promptly meet and confer regarding such responsive ESI. Nothing in this agreement alters or expands the obligations of the Requesting Party or the Producing Party pursuant to Federal Rule of Civil Procedure 26(b).

## VII. FAMILY RELATIONSHIPS OF ELECTRONIC FILES

**A. Family Relationships.** Parent-child relationships between ESI (e.g., the association between an attachment and its parent e-mail, or a spreadsheet embedded within a word

processing document), must be preserved by assigning sequential Bates numbers to all files within a parent-child group, and by providing accurate attachment ranges for those files in the Metadata fields required. If a Party withholds as privileged a member of a parent-child group (e.g., a document or portion of a document) that has other non-privileged or privilege redacted members, any parent-child relationship must be similarly maintained by indicating the absence of that document (or portion thereof) using a Bates-numbered slip sheet (or a clearly marked redaction) and the grounds for withholding.

- In the event a document or portion of a document that is part of a family is initially withheld as privileged, but later produced, the entire family must be reproduced with it, along with all required Metadata.

**B. Non-Responsive Attachments.** The Parties agree that “non-responsive” shall generally not be a cognizable basis for withholding documents from production, if a document is associated, through familial relationships, with a responsive document. An otherwise non-responsive attachment to a responsive parent document—or an otherwise non-responsive parent document that has responsive attachments—may not be withheld simply because the document, if it were not in the same family as responsive material, would not be responsive. The Parties agree, however, that the general provisions of this paragraph shall not prejudice any Party’s right to seek exceptions if it identifies particular non-responsive documents, or categories or portions of documents, that it believes are appropriate to withhold or redact on the basis of non-responsiveness notwithstanding their association with a responsive document. In such case, the Parties shall meet and confer regarding the exception sought and may submit any such issue to the Court if agreement cannot be reached.

**C. Microsoft “Auto” Feature and Macros.** Microsoft Excel (.xls) and Microsoft PowerPoint (.ppt) documents should be analyzed for the “auto” features, where documents have an automatically updated date and time in the document, file names, file paths, or similar information that when processed would be inaccurate for how the document was

1 used in the ordinary course of business. If “auto date,” “auto file name,” “auto file path,”  
2 or similar features are identified, the produced document shall be branded with the words  
3 “Auto Date,” “Auto File Name,” “Auto File Path” or similar words that describe the  
4 “auto” feature. Similarly, if a document contains a “macro,” the document shall be  
5 branded with the word “Macro.”

6 **D. Embedded Documents.** Embedded ESI documents (e.g., a spreadsheet embedded within  
7 a word processing document or an audio file in a PowerPoint presentation) will be  
8 extracted, produced as independent document records and related back to the respective  
9 top level parent document (e.g., standalone file, email message, etc.) via the BegAttach  
10 and EndAttach fields referenced in Exhibit 1. The embedded document will be marked  
11 with a “YES” in the load file under the IsEmbedded field referenced in Exhibit 1. Related  
12 embedded documents will be produced within a continuous Bates range.

13 **E. Zero-byte Files:** The Parties may, but are not required to, filter out stand-alone files  
14 identified as zero-bytes in size that do not contain responsive file links or file names. If the  
15 Requesting Party in good faith believes that a zero-byte file was withheld from production  
16 and contains information responsive to a request for production, the Requesting Party may  
17 request that the Producing Party produce the zero-byte file. The Requesting Party may  
18 provide a Bates number to the Producing Party of any document that suggests a zero-byte  
19 file was withheld from production and contains information responsive to a request for  
20 production.

21 **F. Hyperlinked Files:** Hyperlinked files such as those indicated by utilizing Microsoft  
22 Office’s “Share Documents Via Link” feature or other third-party document sharing  
23 platforms will be produced upon reasonable request. Other links in email to non-file  
24 sources, such as a website or QR code destination need not be produced.

25 **G. Compressed Files.** Compression file types (e.g., .CAB, .GZ, .RAR, .ZIP), shall be  
26 decompressed to ensure that a compressed file within a compressed file are decompressed  
27 into the lowest possible compression resulting in individual folders and/or files.  
28

## VIII. STRUCTURED DATA

When responsive information is located in a database—including for enterprise resource planning or human resource applications or databases like Workday, SuccessFactor, Access, Oracle, SQL Server data, SAP, etc.—or other structured or aggregated data source, the Parties will meet and confer to address the production format of responsive data.

## IX. PARAMETERS FOR CUSTODIAL ESI COLLECTION

For each of a Producing Party's current or former employees identified as a Document Custodian, the Producing Party will do the following:

- A. Email.** Collect all emails sent to or received by a custodian regardless of whether such emails are in the custodian's actual email account.
- B. Social Media.** Collect all custodian social media information, including messages, posts, direct messages, emails, video, and text messages;
- C. Cell Phone Data.** Data from current or previous cell phone devices from January 1, 2018 to the present will be preserved by direct device collection using an industry standard collection tool for each custodian. Prior to any culling of cellphone data in response to any discovery requests, the Producing Party will provide for each custodian from January 1, 2018 to the present: (a) a list of cell phone number(s) used by the custodian, including any phone numbers associated with temporary or pay-by-month cell phones; (b) the name of the phone carrier that provided service for each identified phone number, as well as the time period each phone carrier provided service; (c) type of phone, including brand and model number, if known; (d) the present physical location of the device and whether the Producing Party has all necessary password or passcode information for accessing the data on the cell phone, if known; (e) the custodian's practices for backing up data stored on the phone (e.g., cloud backup of an iPhone); and (f) a list of installed communications-related applications on the custodian's phone, including ephemeral messaging applications (e.g., Snapchat), Facebook Messenger, Microsoft Teams, Slack, Wechat, WhatsApp, and other such applications.

- A Producing Party will review the following sources of information on a cellphone used for work purposes, to the extent reasonably available, to identify unique, responsive, and discoverable information; the Parties will discuss and will not unreasonably oppose methods of culling the below sources for responsive information:

- i. **Cellphone Call and Voicemail Logs:** The logs any calls made/received and voicemails left on a cellphone that the document custodian used for work purposes, if any, if the cellphone is in the possession, custody, or control of a Producing Party.
- ii. **Text Messages:** All text messages and/or iMessages on the cellphone device used for work purposes or contained in available backups/archives associated with the device, if any, inf the cellphone is in the possession, custody, or control of a Producing Party.
- iii. **Social Media:** All social media used for work purposes or contained in available backups/archives associated with the device, if any, inf the cellphone is in the possession, custody, or control of a Producing Party.

**D. Phone Records.** Producing Party will either obtain billing statements for all phone number used by custodians or provide signed written authorizations for the release of such phone records form the applicable phone carriers (including facilitating the receipt of such authorizations on behalf of their employees where applicable).

**E. Personal Computer-Based ESI.** Provide a list of every personal computer used by the custodian from 2018 to the present, the model number of the personal computer, the present physical location of the computer, any backups that exist of the information from such a personal computer, and if a personal computer used by a custodian has been lost or destroyed, provide the date upon which it was lost or destroyed, if known. Producing Party will also identify whether any backups for other devices such as an iTunes backup of an iPhone exist on the personal computer.



- 1       **F.     Tablets.** The Producing Party will identify any tablet computers used by custodians  
2       between 2018 and the present.
- 3       **G.     “Contacts.”** A custodian’s contacts (e.g., Outlook contacts or cell-phone based contacts)  
4       will be exported to Excel (or other spreadsheet format agreed upon by the Parties) from  
5       the custodian’s email account and/or cell phone with all Metadata fields included in the  
6       extraction. If a Party desires to redact certain personal information (e.g., phone numbers  
7       for a custodian’s children) then the Parties will meet and confer regarding the categories of  
8       such information to be redacted and how such information will be redacted/excluded from  
9       the responsive spreadsheet.
- 10      **H.     Cloud-Based ESI.** To the extent a custodian used any cloud-based document storage  
11      services (e.g., Dropbox) for work-related purposes, the Producing Party will obtain such  
12      information and produce responsive information. To the extent a custodian backed up a  
13      cell phone or table to a cloud-based storage system (e.g., iCloud), the Producing Party will  
14      obtain such information and produce responsive information. If the Producing Party is  
15      unable to obtain cloud-based ESI for a custodian, then the Producing Party will provide  
16      the Requesting Party an explanation in writing of the circumstances of the inability  
17      including (a) the name of the custodian; (b) the name of the cloud-based system upon  
18      which the data is believed to be stored; (c) a description of the type of ESI believed to  
19      stored in the cloud-based system; (d) the efforts the Producing Party undertook to obtain  
20      the information; (e) whether any data on the cloud-based system is believed to have been  
21      deleted or lost, and, if so, when; and (f) whether or not the user name and password for  
22      such information is known to the custodian or the Producing Party.
- 23      **I.     Enterprise Messaging and Other Communication Tools.** The parties will disclose at the  
24      outset any use of enterprise messaging tools meant to replace email, such as Slack or MS  
25      Teams, including the type of subscription and retention settings in place. The parties will  
26      meet and confer regarding issues relating the parameters for an inclusive collection of  
27      such data, the format of the production, and other unique issues to this type of data.  
28

**J. Social Media Data.** If a document custodian confirms that he or she (1) used Social Media and (2) used that Social Media to communicate with an employee of another Defendant to post or otherwise communicate regarding a subject relevant to the litigation and included within a Request for Production, subject to objections to that Request, then the requested communication(s) must be produced if it is reasonably accessible, in the producing Party's possession, custody, or control, and not withheld as privileged and/or as illegal to produce under applicable privacy laws. The Parties shall meet and confer to the extent there are any issues with respect to the format of such Social Media data.

**K. Administrative Assistants.** An agreement to produce documents for a particular custodian shall include the collection of that custodian's current and/or former administrative assistants.

## **X. PRODUCTION MEDIA**

Documents shall be produced via secure FTP site, rather than through physical media (e.g., CD-ROM, DVD, or external hard drive), unless such transmission is impracticable or the Parties hereafter agree upon on a different method of transmission. Each production shall be accompanied by a cover letter, to be sent via email at the time of production, including the following information:

**A.** Name of the Litigation and its case number;

**B.** Name of the Producing Party;

**C.** Date of the production (mm/dd/yyyy);

**D.** Bates number range;

Confidentiality Designation; and

**E.** Notes regarding any irregularities in the production (e.g., whether it is a replacement production (see below)).

**F.** Any replacement production or Documents shall cross-reference the original production, clearly identify that it is a replacement, and cross-reference the Bates number range that is being replaced. Producing Parties may encrypt their productions and, if so, shall provide a key to decrypt the production in a separate communication via email.

**XI. NON-PARTY SUBPOENAS AND PRODUCTIONS**

Any Party that issues a non-party subpoena shall timely notify other Parties when it receives non-party productions of documents, ESI or other responses from the non-party, and shall provide copies of all such productions in the format in which they were received from the non-party to the other Parties. If the non-party produces documents or ESI that are not Bates labeled, the Receiving Party will Bates-label the documents or ESI before producing a copy to the other Parties if the produced documents or ESI are of the type so requiring consistent with other provisions of this Order.

Nothing in this Protocol precludes a Receiving Party from reviewing a non-party production of documents or ESI and applying (if the non-party produced documents without a confidentiality designation) or elevating a confidentiality designation (e.g., elevating documents designated as CONFIDENTIAL to ATTORNEYS' EYES ONLY) if the production contains highly sensitive information as set forth in the Protective Order.

**XII. OBJECTIONS TO ESI PRODUCTION**

If any formatting requirements or other specifications agreed to in this Protocol are later determined by the Producing Party to be not feasible, or unduly burdensome or unreasonably costly, the Producing Party shall notify the Requesting Party, and the Parties will meet and confer in order to attempt to agree upon appropriate modifications to the Protocol before presenting any such issue to the Court. Notwithstanding anything contained herein to the contrary, a Producing Party shall not intentionally produce ESI in a format not provided for in this Protocol, and if a mistake is made, shall be allowed to remediate that mistake. If an issue arises, the issue may be submitted to the Court after reasonable efforts by the Parties to reach agreement by the meet and confer process.

**XIII. ESI PROTECTED FROM DISCOVERY OR PUBLIC DISCLOSURE**

**A. Privilege Logs.** For any Document withheld in whole or in part on the basis of privilege, work product doctrine or other protection, the Party or third party withholding the Document(s) shall produce a privilege log or logs concerning any information that has been redacted or withheld in whole or in part from that production. The withholding

party's log shall be in Excel format that permits electronic sorting and searching. The privilege log shall include:

- A unique number for each entry on the log.
- Date of document for all ESI and to the extent known or discernible for all hard copy documents. For emails this should be the sent date of the document and for loose ESI this should be the last-modified or create date of the document.
- Custodian of the document. For emails this should be populated with the metadata extracted from the Custodian(s) field. For loose ESI, this should be populated with the Custodian value assigned to the instance of the document in question.
- Author of the document. For emails this should be populated with the metadata extracted from the "Email From" field associated with the file. For loose ESI, this should be populated with the metadata extracted from the "Author" field. For hard copy documents, this will be populated with the "Custodian."
- Recipient(s) of the document where reasonably ascertainable. For emails this will be populated with the metadata extracted from the "Email To" field associated with the file. Separate columns will be included for the metadata extracted from the "Email CC" and "Email BCC" fields, where populated.
- The type of privilege being asserted, for example: (a) AC for Attorney/Client, (b) WP for Attorney Work Product, (c) CI for Common Interest.

#### **XIV. RESERVATION OF RIGHTS**

By entering this Protocol, a Party does not waive any objections it has to requests for documents. Nor does a Party give up its right to review its documents for privilege or any other reason (including to identify non-responsive documents), and the existence of this Protocol cannot be used to compel a party to produce documents without review. The existence of this Protocol does not waive any Party's objections based on proportionality or burden nor does this Protocol otherwise supersede Federal Rule of Civil Procedure 26(b).

1 IT IS SO STIPULATED, THROUGH COUNSEL OF RECORD.

2 **PURSUANT TO STIPULATION, IT IS SO ORDERED.**

3  
4 Dated:

5 \_\_\_\_\_  
6 Honorable Jon S. Tigar  
7 United States District Judge

8  
9 Dated: October 27, 2023

10 By:           [NOT SIGNED]            
11 Joseph R. Saveri

12 Joseph R. Saveri (State Bar No. 130064)  
13 Steven N. Williams (State Bar No. 175489)  
14 Cadio Zirpoli (State Bar No. 179108)  
15 Elissa A. Buchanan (State Bar No. 249996)  
16 Travis Manfredi (State Bar No. 281779)  
17 **JOSEPH SAVERI LAW FIRM, LLP**  
18 601 California Street, Suite 1000  
19 San Francisco, California 94108  
20 Telephone: (415) 500-6800  
21 Facsimile: (415) 395-9940  
22 Email: jsaveri@saverilawfirm.com  
23 swilliams@saverilawfirm.com  
24 czirpoli@saverilawfirm.com  
25 eabuchanan@saverilawfirm.com  
26 tmanfredi@saverilawfirm.com

27 *Counsel for Plaintiffs and the Proposed Class*  
28

1 Dated: October 27, 2023

By:           /NOT SIGNED/            
Annette L. Hurst

2  
3 Annette L. Hurst (State Bar No. 148738)  
4 **ORRICK, HERRINGTON & SUTCLIFFE LLP**  
5 The Orrick Building  
6 405 Howard Street  
7 San Francisco, CA 94105  
8 Telephone: 415 773 5700  
9 Facsimile: 415-773-5759  
10 Email: ahurst@orrick.com

11 *Counsel for GitHub, Inc. and Microsoft Corporation*

12 Dated: October 27, 2023

By:           /NOT SIGNED/            
Joseph C. Gratz

13 Joseph C. Gratz (State Bar No. 240676)  
14 **MORRISON & FOERSTER LLP**  
15 425 Market Street  
16 San Francisco, CA 94105  
17 Telephone: 415-268-7000  
18 Facsimile: 415-268-7522  
19 Email: JGratz@mfo.com

20 *Counsel for OpenAI, Inc.; OpenAI, L.P.; OpenAI GP,*  
21 *L.L.C.; OpenAI Startup Fund GP I, L.L.C.; OpenAI*  
22 *Startup Fund I, L.P.; and OpenAI Startup Fund*  
23 *Management, LLC*

**EXHIBIT 1****FIELDS and METADATA TO BE PRODUCED**

1. The load files accompanying scanned paper documents will include the following objective coding fields, to the extent applicable. To the extent that metadata does not exist or is not reasonably accessible or available for any documents produced, nothing in this Order shall require any party to extract, capture, collect or produce such metadata:

<i>Field</i>	<i>Data Type</i>	<i>Paper</i>	<i>Native Files &amp; Email Attachments</i>	<i>Email</i>	<i>Calendar Entries</i>
BegDoc	Integer - Text	Starting Bates #	Starting Bates #	Starting Bates #	Starting Bates #
EndDoc	Integer - Text	Ending Bates #	Ending Bates #	Ending Bates #	Ending Bates #
BegAttach	Integer - Text	Starting bates # of document family	Starting bates # of document family	Starting bates # of document family	Starting bates # of document family
EndAttach	Integer - Text	Ending bates # of document family	Ending bates # of document family	Ending bates # of document family	Ending bates # of document family
CountAttach	Integer – Text	Number of attachments	Number of attachments	Number of attachments	Number of attachments
Custodians	Text – paragraph Separate entries with “,”		All names of people the document was collected from even if removed from production as a duplicate	All names of people the document was collected from even if removed from production as a duplicate	All names of people the document was collected from even if removed from production as a duplicate
From	Text - paragraph			Sender of message	Sender of Calendar invite
To	Text – paragraph Separate entries with “,”			Recipients of message	Recipient of calendar invite
CC	Text – paragraph Separate entries with “,”			Copied recipients	

<i>Field</i>	<i>Data Type</i>	<i>Paper</i>	<i>Native Files &amp; Email Attachments</i>	<i>Email</i>	<i>Calendar Entries</i>
BCC	Text – paragraph Separate entries with “;”			Blind copied recipients	
Subject	Text - paragraph			Subject of message	Subject of calendar invite
DateSent	Date (mm/dd/yy yy)			Date message sent	Date calendar invite sent (if any)
DateReceived	Date (mm/dd/yy yy)			Date message received	Date calendar invite received (if any)
TimeSent	Time (hh:mm:ss)			Time email sent	Time calendar invite sent (if any)
TimeReceived	Time (hh:mm:ss)			Time email received	Time calendar invite received (if any)
Conversation thread	Text- paragraph		Field of metadata tying related emails, where available in email metadata.	Field of metadata tying related emails, where available in email metadata.	
FileName	Text - paragraph		Name of original file including extension	Name of original file including extension	
FileExtension	Text		Extension of original file	Extension of original file	Extension of original calendar entry
ItemType	Text - paragraph	Identifies if the file is an email, attachment, or stand- alone file.	Identifies if the file is an email, attachment, or stand-alone efile.	Identifies if the file is an email, attachment, or stand-alone efile.	Identifies if the file is an email, attachment, or stand-alone efile.
DateCreated	Date/Time (mm/dd/yy yy)		Date file was created		
DateModified	Date/Time (mm/dd/yy yy)		Date file was last modified		
Author	Text - paragraph		Document author from metadata		



<i>Field</i>	<i>Data Type</i>	<i>Paper</i>	<i>Native Files &amp; Email Attachments</i>	<i>Email</i>	<i>Calendar Entries</i>
Confidentiality	Single choice	Designation branded onto the document	Designation branded onto the document	Designation branded onto the document	Designation branded onto the document
Redaction	Integer - text	Indicates whether a redaction has been made to any page of the document	Indicates whether a redaction has been made to any page of the document	Indicates whether a redaction has been made to any page of the document	Indicates whether a redaction has been made to any page of the document
File/folder path	Text – paragraph	Full path to the file(s) at its original location(s), including for duplicates	Full path to the file(s) at its original location(s), including for duplicates	Full path to the file(s) at its original location(s), including for duplicates	Full path to the file(s) at its original location(s), including for duplicates
FullText	Link	Link to the extracted or OCR text file	Link to the extracted or OCR text file	Link to the extracted or OCR text file	Link to the extracted or OCR text file
DocLink	Link	Link to the native file	Link to the native file	Link to the native file	Link to the native file
MD-5 Hash Value or SHA-1 Hash Value	Integer - Text		Hash value	Hash value	Hash value
Time Zone Processed	Text - paragraph	UTC (Universal Coordinated Time)	UTC (Universal Coordinated Time)	UTC (Universal Coordinated Time)	UTC (Universal Coordinated Time)

**EXHIBIT 2****REQUESTED LOAD FILE FORMAT FOR DOCUMENTS AND ESI****1. Delimited Text File.**

A delimited text file (.DAT File) containing the fields listed in Exhibit 1 should be provided. The delimiters for the file should be Concordance defaults accordingly:

Comma	ASCII character 20 (¶)
Quote	ASCII character 254 (p)
Newline	ASCII character 174 (®)

þBegBatesþþEndBatesþþBegAttachþþEndAttachþþPageCountþþCustodianþ

- The first record should contain the field names in the order of the data.
- All date fields should be produced in mm/dd/yyyy format.
- Use carriage-return line-feed to indicate the start of the next record.
- Load files should not span across media (e.g., CDs, DVDs, Hard Drives, etc.); a separate volume should be created for each piece of media delivered.
- The name of the data load file should mirror the name of the delivery volume, and should have a .DAT extension (i.e., ABC001.DAT).
- The volume names should be consecutive (i.e., ABC001, ABC002, et. seq.).
- If Foreign Language / Unicode text exists, DAT file shall be in appropriate UTF-8 or UTF-16 format.

**2. Image Cross-Reference File (.opt).** The Opticon cross-reference file is a comma delimited file consisting of six fields per line. There must be a line in the cross-reference file for every image in the database. The format for the file is as follows:

ImageID,VolumeLabel,ImageFilePath,DocumentBreak,FolderBreak,BoxBreak,PageCount.

Further definitions and instructions as follows:

ImageID	The unique designation that Concordance and Opticon use to identify an image. This should be the BegBates Number of the Document.
VolumeLabel	The name of the volume.
ImageFilePath	The full path to the image file.
DocumentBreak	If this field contains the letter "Y," then this is the first page of a Document. If this field is blank, then this page is not the first page of a Document.
FolderBreak	Leave empty.
BoxBreak	Leave empty.
PageCount	Number of pages in the Document.

- The name of the image load file should mirror the name of the delivery volume, and should have the appropriate extension (e.g., ABC001.LFP).
- The volume names should be consecutive (i.e., ABC001, ABC002, et. seq.).
- There should be one row in the load file per TIFF image.
- Every image in the delivery volume should be contained in the image load file.
- The image key should be named the same as the Bates number of the page.
- Load files should not span across media (e.g., CDs, DVDs, Hard Drives, Etc.), i.e., a separate volume should be created for each piece of media delivered.

### 3. Sample Data

PROD00000001,VOL001,E:\100\ PROD00000001.TIF,Y,,2

PROD00000002,VOL001,E:\100\ MT00000002.TIF,,,,

PROD00000003,VOL001,E:\100\ MT00000003.TIF,Y,,,4

PROD00000004,VOL001,E:\100\ MT00000004.TIF,,,,

PROD00000005,VOL001,E:\100\ MT00000005.TIF,,,,

PROD00000006,VOL001,E:\100\ MT00000006.TIF,,,,

Summation DII File:

@C HASIMAGE YES

@T EPRG036023 2

1 @DD:\FILES\PRODUCTION\CD0004\VOL0011\Images

2 EPRG036023.tif

3 EPRG036024.tif

4 **4. OCR / Extracted Text Files**

5 OCR or Extracted Text files shall be provided in a separate \OCR\ directory containing  
6 Document level text files.

7 If Foreign Language / Unicode text exists, TEXT files shall be in appropriate UTF-8 or UTF-16  
8 format.

**SIGNATURE ATTESTATION**

Pursuant to Local Rule 5-1(h)(3), I attest under penalty of perjury that concurrence in the filing of this document has been obtained from any other signatory to this document.

Dated: October 27, 2023

By: *[NOT SIGNED]*  
Joseph R. Saveri